Soil data process question -- fine to ...

fan 25 posts since

Apr 16, 2008

Hi,

I have a question about processing soil data, to 12.5km in my case. In the data_proc codes, for example, domdata_from_1km.F90, the output fields are filled with 0.0. However, in the original dataset such as FAO soil color data at 1KM resolution, there are no zero values and the missing values are set as -9999. If that is running ok and is right at 1km,then changing all the missing data point to 0.0 for coarser resolutions may have problems. This is also seen in other processing codes.

Could anyone please clarify what is expected (0. or -9999.) by the land surface models from input datasets?

Thanks!

Xingang Tags: soil, data

sujay *118 posts since*
Sep 20, 2007 1. **Re: Soil data process question -- fine to coarse resolution** Apr 18, 2008 9:17 AM

Xingang,

The data processing programs are not designed to work in grids that do not map exactly to 1km subset domains. So you can't use them to create datasets at 12.5km. Please contact the NLDAS group at GSFC as they have processed these datasets at 12.5KM.

I believe these programs give you an option to specify the desired undef value (optional)

domdata_from_1km out-res(km) input_file output_file input_mask_f undef_v

-Sujay

fan *25 posts since*
Apr 16, 2008 2. **Re: Soil data process question -- fine to coarse resolution** Apr 18, 2008 12:00 PM

in response to: sujay

Sujay:

Thanks! We (Yan and I) have found the limit of the original code. I have modified the codes so that they accept real (floating point) numbers for resolution. I have done this for all the data types. I am attaching one of them here as an example and hope it be helpful for others too.

I haven't dug into the LSMs too much. So it will be very helpful if you or someone else know what number is accepted by the LIS or LSM as a proper undefined value or it doesn't matter to specify any value as undef. Since the processed data is input to model, and the model has to deal with and treat the undef value in a differnt way than valid values.

I found in my case, there was no such thing happened where (mask.gt.0.5 .AND. out(i,j).eq.undef), thus, no filling in was actually done for the soil color data (haven't checked other data types yet).

Xingang **Attachments:**

- domdata_from_1km.F90 (7.4 K)

sujay *118 posts since*
Sep 20, 2007 3. **Re: Soil data process question -- fine to coarse resolution** Apr 18, 2008 12:14 PM

in response to: fan Xingang,

I am afraid changing the resolution variable to a real number is not the right fix, though it will let you run the program. The domdata* program selects the dominant type from the native 1km file and when you have a grid (12.5KM) that does not map exactly to the 1km pixels, special logic need to be implemented to modify the sampling.

LIS does not impose a particular undefined value. The lis.config file lets you define one and it is the user's responsibility to make sure all parameter data use this value.

-Sujay

fan *25 posts since*
Apr 16, 2008 4. **Re: Soil data process question -- fine to coarse resolution** Apr 18, 2008 12:58 PM

in response to: sujay

Sujay:

Thanks for clarifying the usage and definition for undef value. I missed the point that it is set in lis.config. I'll pay attention to the situation when a fill-in value do happen.

To get 12.5km dominant values, my current codes count a 1km pixel when half of it falls into the big 12.5km cell. For example, big cell No.1 counts 1km cells (1:13,1:13), No.2 counts 1km cells (13:25,1:13), No.3 counts (26:38,1:13), No.4 counts (38:50,1:13), and so on. This way, rows/cols of 13, 38, 63, 88, ... in the naticve 1km grid are counted twice, and this may only cause errors of less than ~8% (=(13/2+13/2)/13**2)*100%); and error happens only when there are very close numer of 1km cells of two categories. Choosing either one category as the dominant has already included a large error (~50% in the dominant value search approach), even for a grid that exactly maps onto whole 1km cells. So I hope this would be tolerable.

In our project, we have to work on a 12.5km grid, which also matches resolution of NLDAS forcing data.

If anyone has a better solution, please share.

Xingang

sujay *118 posts since*
Sep 20, 2007 5. **Re: Soil data process question -- fine to coarse resolution** Apr 18, 2008 1:08 PM

in response to: fan Xingang,

I guess you approach is okay. As I said earlier, the NLDAS folks are already done some benchmarks on a 12.5KM grid using the LIS code. So if your project requires any sort of cross-comparisons with prior NLDAS studies, I recommend you compare the parameter datasets that you derive with the ones they use.

Thanks,

-S

fan *25 posts since*
Apr 16, 2008 6. **Re: Soil data process question -- fine to coarse resolution** Apr 18, 2008 1:16 PM

in response to: sujay

Sujay:

Thank you for the information and the suggestions, and also for your time talking about the issues.

Xingang